THE EFFECTS OF PROBLEM-BASED LEARNING ON NATIONAL BOARD SCORES, CLINICAL EVALUATIONS AND RESIDENCY SELECTION OF MEDICAL STUDENTS

A Dissertation

Presented to

The School of Education

Drake University

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Education

---

by

Rodney Louis Tomczak

November 1991

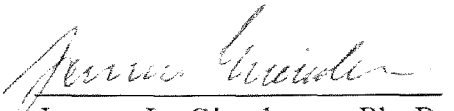
THE EFFECTS OF PROBLEM BASED LEARNING ON NATIONAL BOARD SCORES, CLINICAL EVALUATIONS AND RESIDENCY SELECTION OF MEDICAL STUDENTS

By

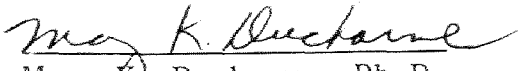
Rodney Louis Tomczak

Approved by the committee:

Charles S. Greenwood, Ph.D.
Associate Dean
Chair



James J. Giordano, Ph.D.
Assistant Professor



Mary K. Ducharme, Ph.D.
Associate Professor

James Romig, Ed.D.
Dean, School of Education

The Effects of Problem-Based Learning on National Board Scores, Clinical Evaluations and Residency Selection of Medical Students

An Abstract of a Dissertation by
Rodney Louis Tomczak
September 1991
Drake University
Advisor: Dr. Charles Greenwood

Problem Statement: There is much support in the literature that medical students are not able to remember facts taught during the first two years of medical school, assimilate those facts to care for patients, or solve patient problems adequately. Some medical educators have tried problem based learning as an alternative to the traditional lecture based curriculum. Few studies have addressed the short term outcomes of National Board Scores, Clinical Evaluations and Residency Selection for students encountering a problem based curriculum.

Procedures: This study compared the Classes of 1990(Traditional Curriculum) and 1991(Problem Based Curriculum) of the University of Osteopathic Medicine and Health Sciences, College of Podiatric Medicine. Those Parameters compared for statistical significance by ANOVA included six sections of the National Boards, Clinical Skills, Knowledge, Interpersonal Qualities and Self-directedness. A Chi-Square Test for Independence was used to compare success in residency selection for the two classes.

Findings: There was no statistical difference on the Orthopaedic, Radiology, Dermatology, Surgery or Podiatric Medicine sections of the National Boards between the Classes of 1990 and 1991. The Class of 1990 scored significantly higher ($p=.036$) on the Community Health section of the National Boards. There was no statistical difference on Clinical Evaluations except for Interpersonal Qualities on which the Class of 1991 scored significantly higher ($p=.003$). In Residency Selections, the Class of 1991 placed in desired residencies statistically better than the Class of 1990 ($p<.05$).

Conclusions: There were significantly better Interpersonal Qualities Evaluations and Residency Placements for the Class of 1991 which encountered a Problem Based Curriculum than the Class of 1990.

Recommendations: Long term outcomes and levels of Self-directedness should be evaluated and compared between physicians who encountered different curriculums while in medical school.

Table of Contents

List of Tables

Chapter 1

Introduction

Background of Medical Education in the United States

Prior to World War II, medical education was almost the same at every medical school in the United States. It was based on the Flexner Report (1910) which had a double mission. The first mission was to aid in the abolishment of the proprietary medical schools which were apprenticeship programs run by individuals. The second mission was the standardization of medical education across America so that a certain amount of unity in the curriculum could be guaranteed. The medical educational system stressed lectures, laboratory sessions and numerous examinations so that the quality of the process could in some way be safeguarded. The first year focused on the normal non-diseased human being with courses in anatomy, physiology, biochemistry, histology and microbiology and were often called the basic sciences. The second year centered around the abnormal and the diseased with courses like pathology, pharmacology, infectious diseases and some of the interventions that medicine had available. The third year signaled the introduction of the student to live patients. Some time was spent in the hospitals, but the majority was still devoted to the didactics of classroom medicine with lectures and examinations that had to be passed. The fourth year of medical school offered more hospital and clinic time, but it was not without its exams (Flexner, 1910). In fact,

it was much like the last nineteen years of school even though these were mature adults.

The five years of World War II offered the most incredible concentration of challenges, changes, stresses, achievements and catastrophes to medicine that one could imagine. Government funds became available for research and for the building of new research facilities; the armed services became involved in the selection of medical students; in places they never thought they would be, young physicians were called upon to meet challenges in medicine and psychiatry that they had never been exposed to in school or training. In addition, the call to active duty left the civilian population without adequate medical care.

Not only did the World War II experience result in an explosion of medical knowledge, but the baby-boomer generation was about to be born and become the Woodstock generation with their own paradigm of health. Believing that the system espoused by Flexner could continue without a need for change, few medical educators had the vision to see what would be needed for this upcoming generation. The paradigm for medical education already existed, and changing it would be as difficult as changing any other paradigm. In the case of medical education, the final convertees to a new paradigm would be those in the highest positions within the profession, the faculties and administrations of the medical schools.

Case Western Reserve University in Cleveland, Ohio had a Board of Directors who felt change was inevitable. They appointed a new

Dean of the Medical School with the charge of revamping the entire medical curriculum. When this occurred, there was absolutely no idea what the finished product would look like, but they knew that medical education as it then existed had to be changed (Williams, 1980).

Up until this time, medicine had failed to do any research in education (Ham, 1964). The dean of Case Western appointed a Committee on Medical Education and a Division of Research in Medical Education (Wearn, 1956). The intial action taken by this Committee on Medical Education was to enlist the aid of professional educators outside Case Western. The Department of Education of the University of Chicago was chosen with John Ginther, Ph.D. serving as the chief evaluator for Case Western Reserve University (Harris, Horrigan, Ginther & Ham, 1962).

During the 1950's Cyril O. Houle was at the University of Chicago along with his student Malcolm Knowles (Knowles,1984). Their focus of study was on the adult learner and the differences between adult and child learners. The effect of Houle and Knowles can be traced through Case Western and its then innovative curriculum to the present day problem-based curriculum medical schools. Their philosophy of adult learners, or as Knowles called it "andragogy" (Knowles, 1980), differed from traditional pedagogy in some basic assumptions. The adult was self-directed and independent, ready and eager to learn, was task or problem centered, naturally curious and motivated. Experience played a large role in the learning process.

The philosophy of the Case Western program had strong roots in the Department of Education of the University of Chicago. Subjects were presented linking basic sciences and clinical sciences from day one thus showing the student the utility of the learning. They were exposed to patients the first week of school to further stress the practicality of the learning. Students were given free time to explore what they felt they had to learn and to educate themselves. Learning in an effective manner and planning the curriculum for the student were emphasized. Effort was given to ascertaining how the student should be treated to encourage individual growth as a person, as a student able to educate himself in knowledge and in the use of knowledge for the problems of medicine (Ham, 1962).

Aside from a change in philosophy, the first real curricular change that emphasized the methods of adult learning was a pilot study in teaching hematology with emphasis on self-education by the students (Harris, Horrigan, Ginther & Ham, 1962). The method used was the assigning of reading material which was to be completed in time for a clinical conference in five days. At the conference a patient with the particular anomaly under study was presented to the whole class. After the presentation, the students went to the laboratory where actual procedures were performed on that patient's blood. The students were then given a problem to work on individually. A brief description of a case was given and the student was asked to commit to ordering laboratory studies and defend the decision. The student turned the page and the results of the tests that were actually done

were provided. The student was then asked to interpret the results and justify the interpretations. Again the student was asked what further studies were indicated and why. Having completed this assignment, the student examined the actual results and was instructed to begin managing the patient.

Students using this method of learning were compared to classes immediately prior to the implementation of the experiment with the assumption that they did not differ. Results showed an upward shift in achievement on the comprehensive examination in hematology for the experimental class and "opinionaires" evaluated by the University of Chicago revealed that the students found the experiment " not unfavorable" (Harris, Horrigan, Ginther & Ham, 1962). The authors issued the caveat that although attempts were made to make this a "tight" study, the results could not be viewed as conclusive until and unless further investigations supported these findings.

Miller (1962) stated that his studies showed that examinations in medical school rarely went beyond the comprehension level of Bloom's Taxonomy. Even then, the concern should not be what the student has at the end of the course but rather what is kept and becomes part of life. He asked sophomores, juniors and seniors to take a second time the examinations they had passed as freshmen. No student passed gross anatomy and only a few histology and physiology. Whether the student had come from the upper or lower quarter of the class made no difference. The members of the clinical faculty were asked to take the same exams. A board certified internist scored a thirty in biochemistry

and an Associate Professor of Surgery got a like score in anatomy. Retention of classroom learning and subsequent transfer to a clinical situation seemed to be a significant problem.

At the same time as Miller's study, Howard Barrows was a clinical neurologist in Canada. He developed simulated neurology patients that were standardized so that faculty evaluations could provide more helpful data to the student (Barrows & Abrahamson, 1964). It revealed that students had a paucity of basic knowledge that they could apply to the patient problem. The findings seemed paradoxical since he was closely associated with the pre-clinical curriculum and was sure that the students had been exposed to and passed courses in neuroanatomy, neurophysiology and clinical neurology.

Many shared this same feeling that students were coming to the clinical years ill prepared (Barrows & Tamblyn, 1980). Some suggested an "inverted curriculum" where students would have two years of patient exposure then two years of basic science. Miller (1978) also showed that the retention of basic science information decreased at the same rate as retention of nonsense syllables. West (1966) summarized the fallacies of this traditional approach by pointing out that logic and research proved that the traditional educational approach was ineffective and inefficient.

Significance of the Problem

McGuire (1972) felt that many medical schools, if they evaluated correctly would discover that their graduates did not possess adequate problem solving skills. The work begun at Case Western Reserve has

ground to a halt. Although Case Western has a curriculum which differs significantly from other medical schools, it does not have a problem-based curriculum (Williams,1980). The results of the problem solving course in hematology have not been reproduced in other courses and other schools. Opponents to change are able to simply state that no evidence exists to show that problem-based learning is any better than the lecture method. The lecture method of instruction has become the gold standard against which all other methods are compared. The fund of knowledge within medicine continues to grow while the time devoted to medical school remains constant. It is an accepted fact that medical students cannot possibly learn everything they need to know while in the undergraduate medical education system. They are, however, responsible both morally and legally for new developments that occur after they graduate. This problem is so significant that every state has mandated that physicians must attend continuing education courses each year to stay abreast of current developments. These courses are lecture based and the physicians sit there passively absorbing a few facts just as they did while in medical school.

The few schools that have adopted problem-based learning have published few reports of outcomes. When this has happened the students have usually been specially selected for the problem-based curriculum or placed in a parallel track with conventional students. Without results that are more generalizable, conservative medical

educators view conclusions concerning problem-based learning in medical schools reluctantly.

Problem Statement

There could be measurable differences in outcome measures between students who have a problem-based curriculum and those students who have not had a problem-based curriculum. Little is actually known about either short term outcomes such as Part II National Board Scores, clinical evaluations or placement into desired hospital residencies when comparing students of a problem-based with those who had a more conventional program. Long term goals of a problem-based curriculum, such as self-directedness in learning can not be examined at this time.

The overriding goal to be achieved by any medical school is the demonstrated ability to identify, analyze and manage clinical problems in order to provide effective, efficient and humane patient care. The enabling objectives consist of a knowledge portion, which has been addressed, a skills portion, which includes critical appraisal and clinical skills and the development of personal qualities. The student is evaluated in simulations, often with difficult patients, not only by the instructor but also by other students and that student himself. This process aids in the development of both skills and personal qualities that are necessary to achieve the overriding goal. Although objective testing has for the most part been used as a primary outcome measure, it fails to take into account these other objectives relating to demonstrated skills and personal growth. It only measures the

knowledge portion of the curriculum. Clinical and interpersonal skills are extremely relevant to the general population but before there can be general acceptance of a problem-based curriculum which strives for more than just knowledge, the program must show that there is no diminution in the amount learned by the students. Generally for assessment, standardized tests such as the National Boards are used.

Attempts have been made to use the National Boards as a predictor of success in subsequent years, but no significant correlations have been demonstrated (Gonnella & Hojat, 1983; Ronai, Golmon, Shanks, Schafer & Brunner, 1984; Veloski, Herman, Gonnella, Zeleznik & Kellow; 1979, Williams, Sachs & Veloski, 1986). Nonetheless, if students that experience a problem-based curriculum do not score well on the National Boards, the program is doomed to failure. For this reason, the results of the National Boards have been chosen as a dependent variable for both classes being studied.

If a problem-based curriculum results in better interpersonal skills for the young student-doctor, this should be reflected in the evaluations that are submitted from physicians acting as trainers in teaching hospitals. A comparison of the evaluations measuring these skills will be compared for the two classes in the study.

The residency selection process is a most important process in the medical school experience (Tomczak & Dowdy, 1988). If a student is not selected by a residency program that is highly desired, the student feels that the medical school experience has been disappointing. Residency selection is based on a number of parameters. Students

spend time at hospitals functioning as externs during their fourth year. The residency director of that hospital has ample time to evaluate the student, but since a student cannot devote quality time to every hospital that is desirable, the selection process is based on an interview which often consists of a patient problem analysis rather than an exercise in recalling medical trivia. Because the problem-based learning experience should enhance not only skills but also interpersonal relationships, it is further hypothesized that students encountering this type of curriculum will place better in residencies of their choice regardless of exposure through an externship or merely the interview process than students who did not have a problem-based curriculum.

Assumptions of the Study

The major assumption of this study is that the classes of 1990 and 1991 of the University of Osteopathic Medicine College of Podiatric Medicine and Surgery are alike in all parameters except the independent variable which in this case would be the type of curriculum which was encountered in the clinical courses. This can be evaluated by examination of college GPA and scores on the Medical College Admission Test. By comparing students from two classes at the same school, the design is tighter than comparing two different schools where members of the faculties would differ.

A second assumption is that the National Board Examination is reliable and valid. The examination is supervised by Educational Testing Service of Princeton, New Jersey and in the past both validity

and reliability have been high. The examination given in 1990 has not been changed for the class of 1991.

General Hypotheses

It is hypothesized that:

1) There is no difference in National Board Part II scores between the Class of 1990 and 1991 of the University of Osteopathic Medicine College of Podiatric Medicine.

2) There is no difference in clinical evaluations of the fourth year hospital clerkships for the classes of 1990 and 1991 of the University of Osteopathic Medicine College of Podiatric Medicine.

3) There is no difference in the success rate of securing desired residencies between the Class of 1990 and the Class of 1991 of the University of Osteopathic Medicine College of Podiatric Medicine.

Definitions

Problem Based Learning: The educational method in which students are presented with the patient's presenting picture in simulation formats that allow for free inquiry. They employ all steps in the clinical reasoning process to establish the data base relative to their hypotheses. Active, teacher-guided exploration and evaluation of the problem, using facilitation skills, directly activates the student's prior knowledge, much of which may be beyond conscious recall, for review and association with new learning. After an episode of self-directed study is completed, the students are asked to evaluate the information resources they used and then to return to the patient problem. They continue to the patient problem in light of the new knowledge and

return to another round of self-directed learning where they continue to analyze and synthesize until the patient problem can be solved (Barrows & Tamblyn,1977).

National Board Examination, Part II: An examination given to medical students in the Spring of their fourth year of school that examines knowledge of the clinical sciences such as medicine and surgery. It is monitored and supervised by Educational Testing Services (ETS) of Princeton, New Jersey.

Clinical Evaluations: A form completed by practicing physicians who train medical students in hospitals that are not on the campus. The physicians are adjunct clinical faculty members and have been instructed in the desired procedure for completing the form. This evaluation form utilizes a 5 point Likert scale to measure knowledge, skills and interpersonal relationships.

Residency Selection: During the fourth year of medical school the students apply to various hospital residency programs in the United States. The hospital residency selection committee reviews grades, clinical evaluations, letters of recommendation and any other information available. A pool of candidates is then usually invited for an interview. After the interviews the hospital ranks the candidates in order of preference. The student, after having interviewed at the hospitals also rank orders the hospitals in order of desired preference. A computer service matches students to hospitals on an assigned date with the match always being in favor of the student. That is, the student will go to the hospital he or she ranks highest that chooses him

or her. The student will not go to a hospital that he or she has not chosen even if the hospital chooses that student as its number one pick.

Chapter 2

Review of the Literature

Philosophical Origins of Problem-Based Learning

Knowles (1984) felt that one of the assumptions of andragogy or adult learning was that the orientation to learning for the adult was problem-centered. Adults devoted time and energy to learn something to the extent that they perceived that it would help them perform tasks or deal with problems that they confront in their life situations. Knowles (1988) gave credit to Eduard Lindeman for launching the philosophical stream of adult education in his The Meaning of Adult Education (1961). In this work, Lindeman laid out four assumptions of andragogy:

1. Education is life, not a preparation for it.
2. Education revolves about non-vocational ideals.
3. The approach to adult education should be via situations, not subjects.
4. The resource of highest value is the learner's experience.

Lindeman was a strong proponent of the Danish system of adult education. He visited Denmark often, claiming it to be the homeland of his parents. Stewart (1987) in his biography of Lindeman stated that the main influence on Lindeman was N.F.S. Grundtvig (1783-1872), a Danish theologian-philosopher-poet who authored a body of writing estimated at 150 tomes, most of which is untranslated and unpublished.

Stewart (1987) discovered that Lindeman was actually of German descent but claimed to be Danish. It appeared that he had a deep psychological need to be something other than German. For this reason he was not able to read any of Grundtvig's original works but wrote much concerning the effects of Grundtvig on the Danish educational system. This seemed to consist of eight philosophical pillars upon which Grundtvig's philosophy of learning was based (Grundtvig, 1976).

Enlightenment of life. This referred to Grundtvig's desire for mankind to lead a useful and pleasureable human life. It was the cornerstone of his philosophy and eventually resulted in his break with the Catholic Church. The Catholic perspective of redemption that espoused that suffering was necessary for redemption was contrary to Grundtvig.

The living word. People were not redeemed by winning God's grace. The only way to lose salvation was to reject it. Hence, less emphasis was needed on rules and regulations but more should be placed of other individuals seeking enlightenment.

Change. The state must be seen as a creation of the citizens and existed for them. The state, however, tried to remain static,which according to Grundtvig was against the natural order of things. In order to insure change, prevent the state from placing itself before the citizens and protect the people, the living word was necessary.

Fellowship and freedom. If one did not have fellowship, one could not be free, and if freedom was not facilitated in self and others,

fellowship could not be enjoyed. This fellowship is accomplished by the living word and led to the enlightened life.

Teacher-student interplay. Bugge (1965) called Grundtvig's theory of teaching "reciprocal-teaching". Students were to teach the teachers rather than the teacher lecturing the student. Lectures were viewed as counter-productive. The teachers were supposed to be supportive, give guidance and serve as models.

Student-student interplay. Because students could not stay in school their entire lives, they had to learn how to become proficient at dialoguing. This process would ensure continued learning outside the classroom.

National poet-educational community interplay. For Grundtvig, poetry was a method of knowing one's self. This study led to a personal interpretation of life that one could comfortably live with. In order to accomplish this, the country's poets visited the schools and offered their expertise to the students so that enlightenment of life could be achieved.

The model. In Grundtvig's school students studied what they agreed on, no real subject matter, no tests, grades or degrees existed. The primary method of instruction was dialogue. The students and teachers lived and worked together while at the school.

The following assumptions concerning adult education can be traced to Knowles and his basic theory (Knowles, 1984).

The need to know. Adults need to know why they need to learn something before they undertake to learn it. This need could be

created by placing students in real or simulated situations where they become aware of what they need to know. The teacher-student interplay and educational model of Grundtvig enhances this possibility.

The learner's self-concept. Adults resented situations where they sense others were imposing their wills on them. In education adults often expected to be taught even though they were self-directed in other aspects of their lives. This psychological problem often led to dropping out of the educational experience. This conflict could be resolved by the supportive guidance of the adult educator and a model that emphasized learning, not grades and tests.

The role of the learner's experience. Adults were a rich source of knowledge based on their experiences and age. For these reasons experiential techniques such as problem-solving, group discussion and simulations should be emphasized rather than lecture techniques.

Readiness to learn. Students often decided what they needed to know based on some life experience or lack of knowledge for that particular situation. Students sought the enlightenment of life that they required for success.

Orientation to learning. Adults devoted time to learning something that they perceived would help them solve the problems in their lives. Hence their orientation was problem-centered, task-centered or life-centered. This philosophy was the enlightenment that man sought when he was presented with a problem not readily solvable.

Motivation. Adults were motivated to learn because they desired increased satisfaction, self-esteem and quality of life. They wanted to

continue growing and developing. They desired the living word and the enlightenment of life.

It was this philosophy of experiential learning, problem-situations and teachers as facilitators of discussion that led to a desire to understand the life situation in which the adult found himself or herself that can find its roots in a theory that preached that the ultimate reason for learning was the enlightenment of life. This enlightenment could only be achieved through the living word and dialogue about experiences rather than through books.

The Psychological Origins of Medical Problem-Based Learning

In 1890 T.C. Chamberlin published a paper in Science entitled The Method of Multiple Working Hypotheses. This seemed to be the first mention of problem solving as a method of learning. Chamberlin thought that if the intellectual method used by the student in learning was the "ruling theory", then a certain intellectual affection would exist. Learning should be dominated by impartial intellectual rectitude. When this did not occur, the learner was buying into a paradigm of the teacher without questioning or being allowed to question.

When one worked with just one hypothesis, the inquiry that was made was done not for the sake of the hypothesis but for the sake of facts. Under the method of the "ruling theory", the stimulus was directed to the finding of facts for the support of the theory. Under the one hypothesis theory, the facts were sought for the purpose of ultimate induction and demonstration, the hypothesis being but a means

for the developments of facts and their relations and the arrangement and preservation of material for the final induction.

To guard against this, Chamberlin recommended the method of multiple working hypotheses. This method was directed against the most radical defect of the other two methods; intellectual parenting. By intellectual parenting Chamberlin meant that once an idea or diagnosis was conceived it becames part of the author and he tried everything to prove that idea or diagnosis correct. All objectivity was lost. By using the multiple working hypotheses method, one entertained every possible explanation of a phenomena and developed every tenable hypothesis.

Chamberlin felt that education was largely a doctrine of what he called "pedagogical uniformitarianism", a synonym for "educational paradigm". Education according to Chamberlin should use the same multiple working hypotheses for teaching and learning that science should use for new discoveries.

Dewey (1933) attempted to define a type of thinking that he called reflective thinking. It involved a state of doubt or perplexity (a problem) and an active search of previous experiences and knowledge for materials that would resove the problem. Akin to Piaget's state of disequilibrium and an essential ingredient in his cognitive theory; the state of doubt must be sustained and protracted and itself become a stimulus to inquiry. It was not impulsive or routine but enabled one to act in a deliberate fashion to attain what was lacking. Dewey (1933)

listed five steps to accomplish this process of reflective thinking. They included;

1. Suggestion of an initial action which was not carried out because of doubt.

2. Intellectualization of the diffuculty into a problem to be solved

3. Hypotheses formation to further guide reasoning

4. Mental elaboration of the idea

5. The testing of the hypothesis

These five steps appeared to be a form of the hypothetical-deductive reasoning process, a form of problem solving or a general heuristic.

Miller (1956) reviewed much of what was known at that time about receiving, processing and remembering information. He felt that there were severe limitations on the amount one could actually put into long term memory at any one time and later recall. It appeared that no matter who performed the study or where it was done, most humans were limited to the magical number seven, plus or minus two bits of information.

Bruner, who was a colleague of Miller at Harvard thought that the major problem with the memory process was not the amount that could be stored, but rather it was the retrieval (Bruner, 1961). The key to retrieval was knowing where to find information and how to get there. Looked at from the retrieval side, the process of memory then became a process of problem solving. Material that was organized in terms of a person's own interests and cognitive structures was material

that had the best chance of being accessible in memory. Therefore, the very activities that characterized "figuring out" or "discovering" things for oneself would seem to have the effect of making material more readily accessible in memory.

This ability to "figure out" or as Bruner called it "learning the heuristics of discovery" probably occured initially through the use of analogies. It was his hypothesis that only through the exercise of problem solving and the efforts of discovery and practice that one was more likely to generalize what one had learned into a style of problem solving or inquiry that served for any of task encountered. Bruner was not sure that the technique could actually be taught as such, but he was convinced that the only way to improve was to engage in the inquiry.

Katona (1940) did not think that the learning and teaching of problem solving skills was that much of an unknown. By stating higher order rules along with the expected goal for the problem, the problems could be solved. More impressive, however, was the use of illustrations to stimulate the learners to discover the rule for themselves.

Gagne (1977) summarized the conditions needed for effective problem solving to occur. They fell into two categories, conditions within the learner and conditions in the learning situation. The learner had to be able to call upon his own previous experiences, specifically, previously learned rules. The learner also had to be able to use the cognitive strategies that were previously learned. These basically

consisted of divergent thought, assimilation of theories and deductive conclusions. The conditions in the learning situation consisted of verbal instructions that may stimulated the recall of relevant experiences. Initial guidance or channeling of thinking was important to inform the problem solver of the goal of the activity.

Problem solving as a method of learning then required that the learners discover certain higher-order rules without specific help. These higher-order rules could be effectively generalized to many situations and were highly resistant to forgetting (Worthen, 1968). The newly learned higher-order rule then became a simpler rule to be used in the discovery of new higher-order rules.

Measurements of Academic Achievement and Clinical Skills in Problem-Based Medical Education in North American Literature

Very little emphasis has been placed on the measurement of student outcomes in institutions that employ a problem-based curriculum. Most of the research work in medical problem-solving has concentrated on three main research prardigms: (a) clinical reasoning from an information processing standpoint; (b) a judgement approach which utilized correlational statistics; and (c) the decision-analysis approach which investigated choice under conditions of uncertainty (Elstein & Bordage, 1979). As a result the number of investigations into the outcomes was limited.

Harris, Horrigan, Ginther and Ham (1962) reported on the pilot study they instituted at Case Western in hematology, the beginning of

problem-based, self-directed education in medicine. The research question they asked was, " Will the Class of 1960 (problem-based) score differently on the hematology comprehensive examination than the Class of 1961 (conventional curriculum)". This quasi-experimental design that used an analysis of variance to investigate the performance of the two classes. The researchers established that the two classes were essentially the same on other academic parameters such as previous performance in other courses. The major threat to the validity of the study was the fact that the comprehensive examinations were not identical, however the tests were analyzed and "determined to be comparable". The experiment utilized the entire classes so that random selection and assignment were not issues. The results of the examinations yielded the following information.

1. The mean score was almost five percentage points higher for the problem-based class.

2. An analysis of variance revealed statistical significance at the .01 alpha level when comparing the two sets of scores.

3. In the problem section of the exam, an analysis of variance revealed the Class of 1961 scored higher at an alpha level of .005.

4. In the non-problem section of the test there was no difference between the two groups.

The fact that this study was limited to one course (hematology) in one year at a single medical school detracted greatly from the generalizability of the conclusion that there was a significant upward

shift in the achievement of students in the Class of 1961. Even the authors issued a caveat concerning generalizations.

Coulson (1983) published the results of a problem-based learning trial from Southern Illinois University as Southern Illinois University began the transition from a conventional curriculum to a problem-based curriculum. The researcher utilized the first year class in two successive years, although the dates were not mentioned. Coulson measured performance on the same multiple choice final exam that measured anatomy, behavioral science biochemistry, pharmacology and physiology. No significant differences between the two classes were found on any of the disciplines examined. Twelve weeks after the examination, the classes were reevaluated and the rate of decline on the test scores was significantly less for the problem-based group.

This study was retrospective in nature and quasi-experimental in design. There was a possibility that if the same examination was given the second year of the trial the students could have had access to it. Perhaps the students in the control group told the experimental group that they were retested on the same subject matter 12 weeks after the first examination and maybe the experimental group might want to "brush up" just in case the same thing happenend again. Again the entire classes were used so randomness was not an issue.

Jones, Bieber, Echt, Scheifly and Ways (1984) from Michigan State University College of Human Medicine repoted to the first international Symposium on Problem-based Learning held in Maastricht, the Netherlands in 1983 about their experiences with a two track

curriculum. After a ten week introductory course that was the same for all, students then selected one of the two tracks. Track one was a conventional curriculum while track two was problem-based. In the nine years that results were reported, 304 out of 915 opted for the problem based curriculum while 611 chose a conventional format. Thirteen of those in the problem-based curriculum dropped out of school while 12 from the conventional curriculum left school. No reason for this was given.

Mock National Boards, called Shelf Exams were given to the students who entered in 1977 and 1978. Overall there was no statistical difference between the two groups on Part One of the Boards, but the problem-based learners scored significantly better ($p<.05$) in the physiology, biochemistry and behavioral sub-sections. Pre and Post Clerkship scores for all students were not statistically different, nor were scores on the FLEX examination or on measures of clinical competence.

Students in this College of Human Medicine were able to self-select the program that they chose to follow. There was basically no difference in the outcomes that were measured, however there was no mention of any counselling provided students, but students reported that they chose their track mostly based on their perceived learning style. Omitted from some of the data were the students who chose the Upper Peninsula program where the Track 2 learning mode was the only one available. No mention as to their mortality or reasons for choices was given.

The University of New Mexico (Kaufman, Mennin, Waterman, Duban, Hansbarger, Silverblatt, Obenshain, Kantrowitz, Becker, Samet & Wiese, 1989) published results of six years experience with a problem-based curriculum. This work included results reported by Baca, Mennin, Kaufman and Moore-West (1987). After admission to the University of New Mexico Medical School, an individual applied for admission to the problem-based curriculum. Twenty individuals were chosen for each class. The two curriculums were taught on the same campus which could foster cross-fertilization of educational methods. Selection bias could be a confounding variable and the authors did not address the methods utilized in the selection of participants for each track.

In their analysis of the two tracks the first comparison was on overall performance on National Boards Part I. The performance of the problem-based group had been lower ($p<.0001$) but performance on Part II had been higher ($p<.01$). In their conclusion the authors stated that curricular innovations could influence academic performance but did not state how or in which direction.

Reference was also made to scores on clinical rotations during the third and fourth years. No difference between the two tracks from 1983-1989 could be found.

Eisenstaedt, Barry and Glanz (1990) conducted a study at Temple University School of Medicine that was akin to the original Case Western study in hematology. Students were chosen from the class randomly and invited to participate in a problem-based hematology course. Those

who participated in this experiment rated independent study as less enjoyable than a conventional curriculum ($p=.01$) but felt it would be less difficult compared to conventional lectures than those who volunteered ($p<.0001$).

All students took an objective test after the course and the problem-based group scored significantly lower than the control group ($p<.001$). Two years later the same test was given, but only 54% of the original two groups took the exam again. The authors did not mention any reasons for the high mortality rate. Utilizing scores for both exams, the researchers concluded that the problem-based learners forgot less than the control group ($p<.001$). Since the control group scored so much higher two years previously, this conclusion could easily be attributed to the effects of regression. In any event the researchers warned readers to be cautious concerning interpretations concerning problem-based learning based on one small segment of the medical school curriculum, much like the Case Western researchers 30 years previously.

McMaster University Medical School has reported the results of its problem-based learning in three articles(Woodward, McCauley & Ridge,1981; Woodward,1984; and Neufeld, Woodward & MacLeod,1989). The 1981 study concentrated on early McMaster graduates. The supervisors of interns in Canadian teaching hospitals were asked to rate the performance of McMaster graduates in terms of the average interns in their respective program. The data showed that 26.1% of the McMaster graduates were rated as performing much better than the

average intern, 38.3% were rated as better than the average intern, 28.7% were rated the same as the average intern and 6.9% as weaker. Ratings of graduates from the other Canadian medical schools showed 10.9% were much better than average, 34.8% were better, 43.5% were rated as average and 10.9% were weaker than the average graduate of a Canadian medical school in the internship year. A significant problem with this study exists due to the fact that the term average is never defined and the subsequent subjectivity and interpretation on the part of each evaluator

This study, done in survey fashion, had a compliance rate of 94%, probably because of government influence. The conclusion was that, in a clinical setting, the interns from McMaster performed better than the average intern from other Canadian medical schools. If McMaster graduates stayed in the Hamilton, Ontario, region where McMaster was located, there could certainly be a question of bias. There was also a problem with the defintion of what "average" signified.

In the 1984 study Woodward found the McMaster graduates scored slightly lower than the Canadian average on the multiple choice part of the Qualifying Examination of the Canadian Medical Council but consistently higher than the national average on the patient management portion of the examination. Although the results were not statistically significant, the researchers felt it safe to say that the McMaster graduates did no worse than other Canadian medical school graduates.

The 1989 article from McMaster followed the graduates from 1972. The authors retrospectively found that McMaster graduates had the same first time pass rate on the Medical Council of Canada Examination as the rest of the Canadian medical schools and had a first time pass rate on specialty oral exams that was higher than the national average. Also reported was the fact that 78.6% of McMaster graduates were able to obtain their first choices of residencies while the rate for all of Canada was 58.9%.

The major problems with this longituidinal retrospective study had to do with the location of residencies. If the graduating class was small but the University Hospital at McMaster was able to accommodate most of the graduates the percentages would be abnormally and falsely inflated compared to other schools.

World Literature Concerning Problem-Based Medical Education

European literature offerd a few more insights into problem-based learning, however it must be remembered that the university system and medical training in general was vastly different from the United States. Claesen and Boshuizen (1985) from Maastricht, The Netherlands compared the Maastricht students to students of other medical schools in The Netherlands. The researchers asked subjects to review a case. Having completed this assignment, they were asked to write down what they remembered as important patient data. In difficult cases the Maastricht students were able to remember more important data than students from other Dutch medical students. The results, however,

were not statistically significant. Since the cases for the examination were chosen by faculty members of the University in Maastricht there could be a significant bias which was transfered to the Maastricht students because of contamination and previous contact with these instructors who selected the cases.

Saunders and his colleagues (1987) administered an 80 item multiple-choice test to final year medical students at the University of Sydney and a comparable group of students at the University of Newcastle, both in Australia. The curriculum at Newcastle was problem-based while the Sydney curriculum was conventional. Two hundred forty-three Sydney and forty-five Newcastle students participated in the study with a participation rate of more than 90% for each school. The mean score for the Sydney students was 71% and 67% for the Newcastle school which had problem-based learning. Although the difference was small, it was statistically significant ($p<.05$). This could be accounted for by the large difference in the numbers of participants. The test only measured one discipline, internal medicine, and long term reliabilty was not established on the locally constructed exam.

Conclusion

The literature in medical education pertaining to problem-based literature was sparse and not well constructed. Considering the nature of problem-based learning in medicine it has been difficult if not impossible to construct true experimental designs. Students self-selected or faculty assigned students to a curriculum based on prior

opinions. With so much at stake, it was highly unlikely a student would arbitrarily submit to an experiment for the sake of educational research.

What has been published has been mostly retrospective studies which offer few strong conclusions. Some literature was not favorable to medical education that was problem-based, but the majority of reports suggested that academic outcomes were about the same after the basic science portion of the curriculum and perhaps slightly better after graduation.

More research in problem-based medical education is needed before any conclusions concerning its effectiveness as a method of learning can be made. Results that show the problem-based students performed as well as conventional students will add evidence that this methodology can be successful in medical school.

Chapter 3

Design of the Study

Purpose of the Study

The explosion of medical knowledge means that it is impossible to teach everything to medical students in the four years traditionally allowed for the curriculum. If students spend time memorizing facts, they have no time for real learning and are actually using a technique that is very superficial, and the facts are quickly forgotten (Newble & Clark, 1986). Medical teachers have been slow in realizing that their primary purpose should not be telling students facts, but in helping them perform tasks and think thoughts they could not have done before. Having accomplished these undertakings is still not sufficient for today's medical teacher. The fund of knowledge in medicine changes so quickly that the student must become a physician who is a life-long learner. This skill must be developed during medical school, but the students should not be expected to become independent learners independently (Herber & Nelson-Herber, 1987).

Throughout a career a physician is expected to solve problems on a daily basis. The skill of problem solving must also be learned and practiced if it is to be both proficient and efficient. Learning within the context of problem solving is not new (Chamberlin, 1890), but its acceptance as an accepted method has been slow. This study will add new knowledge to the problem-based curriculum studies by examining a

class of students who were not specifically chosen for a problem-based curriculum on such important outcome criteria as residency selection, National Board scores and clinical evaluations.

Research Design

Population

The two groups, the class of 1990 and the class of 1991 of the University of Osteopathic Medicine College of Podiatric Medicine and Surgery exhibited the following profiles. The Class of 1990 had 62 graduates, 49 men and 13 women. The Class of 1991 projects 57 graduates, 45 men and 12 women. The median age for men in the Class of 1990 was 24.5 years and for women it was 26.0 years. The median ages for the Class of 1991 was 24.8 years for men and 23 years for women. Cumulative undergraduate grade point averages for the Class of 1990 was 2.87 and 2.90 for the Class of 1991. Medical College Admission Test score means for the Class of 1990 and the Class of 1991 are shown in Table 1.

Table One

Medical College Admission Test Score Means

	CLASS OF 1990	CLASS OF 1991
Biology	6	7
Chemistry	6	6
Physics	6	6
Problems	6	6
Reading	6	5
Quantitative	6	5

The Class of 1990 had students from 28 states while the Class of 1991 represented 29 states. Each of the two Classes had one student from Puerto Rico.

Instrumentation

Content validity of the National Boards, which means that the Boards actually measure what they purport to measure, can be assured by the construct of the examination. Professors in each of the disciplines submit test questions. These professors then meet in Princeton where the questions are reviewed by that same group that submitted the questions. Each question must fit within the specific content universe that was previously defined. If one of the group disagrees, the question is deleted. If a question is determined to be too difficult for entry-level physicians by even one participant, it is also deleted.

Face validity is achieved by another group of physicians who review the test. They are young physicians who have been in practice less than five years.

Educational Testing Service utilizes point biserial correlation to assure reliability of the National Board Examination. It is their policy to delete test questions from the exam that have a low point biserial correlation and to repeat those questions that appear to perform well.

Although scores are reported as "Pass" or "Fail", raw scores for the Classes of 1990 and 1991 will be supplied because of the need for continuous data to satisfy an assumption of ANOVA.

Clinical evaluation forms were constructed by the researcher and measure the knowledge, skills and interpersonal objectives set forth in the goals statement of the McMaster University M.D. program (Neufeld,

Woodward & MacLeod, 1989). The form is constructed on a 5 point Likert scale with a score of 5 denoting superior performance.

The final instrument, to determine which ranked hospital the students matched and hence where each student will be doing a residence is a survey form that asks where the student will be doing a residency at and in what position the student ranked that hospital.

Data Collection

Data for the Class of 1990 has been secured at this time. For the Class of 1991, National Board scores were released in August of 1991, clinical evaluations are being collected at present and the national residency match will occured March 12, 1991.

Analysis of Data

The dependent variables of National Board scores, academic average and clinical evaluations should fulfil the assumptions of the ANOVA, namely:

1. normality of distribution
2. equal variance
3. equal interval data
4. independence of scores

These variables will then be utilized to evaluate the effect of the independent variable, the curriculum:

1. conventional, didactic
2. problem-based

Success in residency placement will be evaluated by using a X^2 Independence Test. Students will indicate matching with their first through fourth choices.

Chapter 4

Analysis of the Data

National Board Scores

There are certain assumptions that must be met in order to use ANOVA to analyze the data. These are; 1. Normal Distribution of Data, 2. Equal Interval Data, 3. Homogeneity of Variance, and 4. Independence of Scores. In order to confirm normal distribution, the skewness of each set of data was run on the SPSS-X program. National Board Score skewness by test is shown in Table 2.

Table 2

National Board Skewness by Test Section

NATIONAL BOARD SECTION	SKEWNESS
Orthopaedics	-.391
Radiology	-.285
Community Health	-.994
Dermatology	-.186
Surgery	-.366
Podiatric Medicine	-.418

The score distributions of Dermatology and Radiology showed small amounts of negative skewness. The scores of Orthopaedics, Community Health, Surgery and Podiatric Medicine showed moderate negative skewness. Negative skewness is desirable for tests designed such as these to protect the public from harm and identify that percentage of candidates who do not demonstrate a minimal level of competence for entry-level practice.

The skewness can be explained by the difficulty of the items that appeared on the test. When a licensure test does not exhibit some negative skewness, the test should be investigated. Positive skewness would indicate a test ceiling that was too high. The distribution represented in these scores was deemed to have been normal.

The data from Educational Testing Service was presented in equal interval fashion.

Homogeneity of Variance of the data was confirmed by Hartley's F-Max test. The results are presented in Table 3.

Table 3

National Board Test Scores Homogeneity of Variance by Section

NATIONAL BOARD SECTION	HARTLEY'S F-MAX
Orthopaedics	1.26
Radiology	1.08
Community Health	1.92
Dermatology	1.33
Surgery	1.00
Podiatric Medicine	1.51

The critical value at the .05 alpha level for N=60 was 1.67. Community Health was the only test section that surpassed this level indicating that the null hypothesis (There is no difference in the variance of the test scores of the two groups.) was rejected. This was due to the single score of 62 which was approximately 4 standard deviatons from the mean of 84. When this score was removed, the F-Max Test result was below the critical value. Because ANOVA is a robust test and because there were almost equal populations in the two groups, the violation of this assumption for this particular section was not deemed to be significant.

Test scores on National Board exams are independent.

Clinical Evaluations

Clinical Evaluations were evaluated statistically by ANOVA. They were tested in the same manner as National Board Scores. Table 4 shows results of skewness of distribution.

Table 4

Clinical Evaluation Skewness by Characteristic

EVALUATION CHARACTERISTIC	SKEWNESS
Overall Evaluation	-.698
Knowledge	-.398
Skills	-.640
Interpersonal Abilities	-.830
Self-Directedness	-.746

There was moderate negative skewness in each of the categories evaluated. Again, moderate negative skewness was deemed desireable to protect the public from incompetent practioners. The skewness was not a threat to the validity of ANOVA.

Data for clinical evaluations was on a 1-5 Likert Scale where 5 was highest.

The Hartley F-Max Test was used to assure Homogeneity of Variance. Results are shown in Table 5.

Table 5

Clinical Evaluation Results Homogeneity of Variance by Characteristic

EVALUATION CHARACTERISTIC	HARTLEY'S F-MAX
Overall Evaluation	1.04
Knowledge	1.21
Skills	1.12
Interpersonal Abilities	1.35
Self-Directedness	2.53

Only in the area of self-directedness did the F-Max surpass the critical value. There was a much smaller variance in the evaluations of the Class of 1991 and was the cause of the high F-Max. Because Anova is a robust test and the populations were almost equal, this was not deemed to be a serious threat to the assumptions of ANOVA.

The evaluation scores were independent of each other.

Residency Selection

The two assumptions of Chi Square that were examined were; 1. Independence of Scores and 2. Nominal Data. Each student's selection into a residency was independent and the data used was a number(1-4) which indicated which ranking of hospital chose the student for a residency.

Hypotheses

It was hypothesized that there would be no difference on National Board Scores between the Classes of 1990 and 1991 for each of the sections. The hypotheses were tested by comparing test scores for each of the sections. The groups were compared relative to the Orthopaedics scores in a univariate analysis of variance (facet by Instruction) in an effort to determine statistical significance. The results of that analysis are shown in Table 6. Because of that result, the null hypothesis was not rejected.

Table 6

ANOVA of Orthopaedics

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	6.355	1	6.355	.245	.622
Inst.	6.355	1	6.355	.245	.622
Explained	6.355	1	6.355	.245	.622
Residual	2799.545	108	25.922		
Total	2805.900	109	25.742		

It was hypothesized that there would be no difference in Radiology scores between the Classes of 1990 and 1991. The classes were compared using an univariate analysis of variance. The results of

that analysis are shown in Table 7. Because of the results, the null hypothesis was not rejected.

Table 7

ANOVA of Radiology

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	60.235	1	60.235	2.311	.131
Inst.	60.235	1	60.235	2.311	.131
Explained	60.235	1	60.235	2.311	.131
Residual	2867.542	110	26.069		
Total	2927.777	111	26.367		

It was hypothesized that there would be no difference in Community Health scores on the National Boards between the Classes of 1990 and 1991. Scores were compared by a univariate analysis of variance. Results are shown in Table 8. Because of the results, the null hypothesis was rejected. The Class of 1990 (Traditional) scored significantly higher than the Class of 1991 (Problem-Based).

Table 8

ANOVA of Community Health

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	122.592	1	122.592	4.507	.036
Inst.	122.592	1	122.592	4.507	.036
Explained	122.592	1	122.592	4.507	.036
Residual	2910.491	107	27.201		
Total	3033.083	108	28.084		

It was hypothesized that there would be no difference in scores on the Dermatology section of the National Boards between the Classes of 1990 and 1991. These scores were compared for statistical significance by utilizing a univariate analysis of variance. The results of that test are shown in Table 9. Because of the results, the null hypothesis could not be rejected.

Table 9

ANOVA of Dermatology

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	3.625	1	3.625	.202	.654
Inst.	3.625	1	3.625	.202	.654
Explained	3.625	1	3.625	.202	.654
Residual	2008.165	112	17.930		
Total	2011.789	113	17.803		

It was hypothesized that there would be no difference on the Surgery section of the National Boards between the Classes of 1990 and 1991. These scores were compared for statistical significance by using a univariate analysis of variance. The results are shown in Table 10. Because of the results, the null hypothesis could not be rejected.

Table 10

ANOVA of Surgery

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	9.981	1	9.981	.267	.606
Inst.	9.981	1	9.981	.267	.606
Explained	9.981	1	9.981	.267	.606
Residual	4187.537	112	37.389		
Total	4197.518	113	37.146		

It was hypothesized that there would be no difference in scores on the Podiatric Medicine section of the National Boards between the Classes of 1990 and 1991. These scores were analyzed for statistical significance by utililyzing a univariate analysis of variance. The results are shown in Table 11. Because of the results, the null hypothesis was not rejected.

Table 11

ANOVA of Podiatric Medicine

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	2.954	1	2.954	.130	.719
Inst.	2.954	1	2.954	.130	.719
Explained	2.954	1	2.954	.130	.719
Residual	2481.442	109	22.766		
Total	2484.396	110	22.585		

It was hypothesized that there would be no difference on overall clinical evaluations performed by outside clinical faculty between the Classes of 1990 and 1991. The overall clinical evaluations were compared for statistical significance by using a univariate analysis of variance. The results are shown in Table 12. Because of the results, the null hypothesis was not rejected.

Table 12

ANOVA of Overall Clinical Evaluations

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	.511	1	.511	2.016	.158
Inst.	.511	1	.511	2.016	.158
Explained	.511	1	.511	2.016	.158
Residual	29.127	115	.253		
Total	29.638	116	.256		

It was hypothesized that there would be no difference between the Classes of 1990 and 1991 in the clinical evaluations by outside clinical faculty of the knowledege levels of these students. These evaluations were compared for statistical significance by using univariate analysis of variance. The results are shown in Table 13. Because of the results, the null hypothesis was not rejected.

Table 13

ANOVA of Knowledege Level Evaluations

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	.036	1	.036	.125	.724
Inst.	.036	1	.036	.125	.724
Explained	.036	1	.036	.125	.724
Residual	33.094	115	.288		
Total	33.130	116	.286		

It was hypothesized that there would be no difference in the evaluation of physical skills between the Classes of 1990 and 1991 as evaluated by outside clinical faculty. These scores were compared for statistical significance by using a univariate analysis of variance. The results are shown in Table 14. Because of the results, the null hypothesis was not rejected.

Table 14

ANOVA of Physical Skills Evaluations

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	.508	1	.508	1.712	.193
Inst.	.508	1	.508	1.712	.193
Explained	.508	1	.508	1,712	.193
Residual	34.103	115	.297		
Total	34.611	116	.298		

It was hypothesized the there would be no difference between the Classes of 1990 and 1991 on the clincal evaluations of interpersonal skills as evaluated by outside clinical faculty. The scores were compared by using a univariate ananlysis of variance. The results are shown in Table 15. Because of the results, the null hypothesis was rejected. The Class of 1991 was evaluated significantly better ($p<.05$) for interpersonal skils.

Table 15

ANOVA of Interpersonal Skills Evaluations

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	2.143	1	2.143	9.474	.003
Inst.	2.143	1	2.143	9.474	.003
Explained	2.143	1	2.143	9.474	.003
Residual	26.014	115	.226		
Total	28.157	116	.243		

It was hypothesized that there would be no difference between the Classes of 1990 and 1991 on the clinical evaluation of self-directedness as evaluated by outside clinical faculty. The scores were compared for statistical significance by using a univaraiate analysis of variance. The results are shown in Table 16. Because of the results, the null hypothesis was not rejected.

Table 16

ANOVA of Self-Directedness Evaluations

Source of Variation	Sum of Squares	Degrees of Freedoom	Mean Square	F	Sig. of F
Main Effects	.222	1	.222	.563	.455
Inst.	.222	1	.222	.563	.455
Explained	.222	1	.222	.563	.455
Residual	45.325	115	.394		
Total	45.547	116	.393		

It was hypothesized that there would be no difference in the residency selection process between the Classes of 1990 and 1991. The results of the selection process was evaluted by a Chi Square Test for Independence. The results are shown in Table 17. Because of the results, the null hypothesis was rejected. The Class of 1991 placed significantly better in their desired residencies than did the Class of 1990. Chi Square=(4, n=119)=13.1 $p<.05$.

Table 17

Chi Square of Residency Selection

Observed Frequencies

	First Choice	Second Choice	Third Choice	Fourth Choice	Totals
1990	26	7	3	26	62
1991	40	3	4	10	57
Totals	66	10	7	36	119

Expected Frequencies

	First Choice	Second Choice	Third Choice	Fourth Choice	Totals
1990	34.39	5.21	3.65	18.76	62
1991	31.61	4.79	3.35	17.24	57
Totals	66	10	7	36	119

Chapter 5

Summary and Conclusions

Overview of the Study

Purpose

Medical students have become overrun with facts, so much so that it is impossible to remember enough material to pass the same exam from one year to the next (Miller, 1962). McGuire (1972) found out that medical school graduates did not posses adequate problem solving skills. Students were coming to their clinical years of medical school ill prepared to treat patients even though they had completed two years of school (Barrows & Tamblyn, 1980). Miller (1978) had shown that the information taught during these first two years, the basic sciences, decreased at the same rate as nonsense syllables. West (1966) had shown that both logic and research had proven that the traditional educational approach was both ineffective and inefficient.

Some medical educators began experimentation in the basic science and clinical components of physician education. Because of the Flexner Report (1910), however, the educational process did not wander far from the traditional model. Outside of McMaster University in Canada and New Mexico, few outcomes of problem-based learning have been assessed. It was the purpose of this dissertation to determine the effectiveness of a new problem-based curriculum at the College of Podiatric Medicine and Surgery at the University of Osteopathic Medicine and Health Sciences.

Methods

This study was designed to measure three broad factors. The first was a comparison of average grades on each of six sections of Part II of the National Boards of Podiatric Medical Examiners taken by the Class of 1990(Traditional Instructional Methodology) and the Class of 1991(Problem-based Learning Methodology). These six sections included Community Health, Dermatology, Surgery, Radiology, Orthopaedics and Podiatric Medicine. The General Medicine section was not included because a perceived irregularity occured at the Des Moines testing center in March of 1991 and a retest was administered. The grades on this retest were not released by Educational Testing Service. These scores were evaluated for statistical significance utilizing a univariate analysis of variance.

The second factor examined was the fourth year clinical evaluations of the Classes of 1990 and 1991 by clinical faculty who were not full-time University employed faculty. These adjunct clinical faculty members practiced in various locations around the United States. They evaluated the students' knowledge base, clinical skills, interpersonal abilities and self-directed learning abilities. These four sections were also combined into an overall clinical evaluation score which was the arithmetic average of the four components. These evaluations were on a 1 to 5 Likert Scale where 1 was the lowest and 5 the highest. These scores were evaluated for statistical significance utilizing a univariate analysis of variance.

The third factor evaluated was the success of each member of the Classes of 1990 and 1991 in attaining the residency training program he or she desired most. This was accomplished by asking each member of

the two classes to indicate whether they had matched with his or her first, second, third, or other choice. These rankings were then evaluated for statistical significance utilizing a Chi Square Test for Independence.

Results Summary

Detailed descriptions of attained results were presented in Chapter 4. Those resulted were summarized as follows.

National Boards, Orthopaedics: There was no difference in scores between the Classes of 1990 and 1991 in Orthopaedics.

National Boards, Radiology: There was no difference in scores between the Classes of 1990 and 1991 in Radiology.

National Boards, Community Health: The Class of 1990 scored significantly better ($p=.036$) on this section of the National Boards.

National Boards, Dermatology: There was no difference in scores between the Classes of 1990 and 1991 in Dermatology.

National Boards, Surgery: There was no difference in scores between the Classes of 1990 and 1991 in Surgery.

National Boards, Podiatric Medicine: There was no difference in scores Between the Classes of 1990 and 1991 in Podiatric Medicine.

Clinical Evaluations, Overall Performance: There was no difference between the Classes of 1990 and 1991 in Overall Clinical Performance.

Clinical Evaluations, Knowledge Level: There was no difference between the Classes of 1990 and 1991 in Knowledge Levels.

Clinical Evaluations, Physical Skills: There was no difference between the Classes of 1990 and 1991 in Physical Skills.

Clinical Evaluations, Interpersonal Skills: The Class of 1991 scored significantly higher ($p=.003$) on Interpersonal Skills than the Class of 1990.

Clinical Evaluations, Self-Directedness: There was no difference between the Classes of 1990 and 1991 on Self-Directedness.

Residency Selections: The Class of 1991 placed members in most desired residencies significantly better ($p<.05$) than did the Class of 1991.

Discussion and Practical Significance

It was hypothesized that the students who served as the pioneers in problem-based learning at the College of Podiatric Medicine, University of Osteopathic Medicine would not score differently on National Boards. In three sections they scored slightly lower, one significantly. A meta-analysis according to Glass, McGaw and Smith (1981) was performed on each section to determine an effect size to aid in determination of practical importance. The effect size for Community Health was -.48. An effect size greater than .33 has been determined by researchers to have practical significance (Borg & Gall, 1989). It seemed logical that a negative effect size should also be investigated. The subject matter covered in this section of the National Boards also includes Jurisprudence which was not taught through the problem-based learning method but rather through a conventional lecture format. This occurred during the problem-based curriculum. Non-mandatory morning lectures were offered. There was poor attendance by students who had been exposed to problem-based learning. Very little Community Health or Jurisprudence was offered through the problem-based method.

The effect size for Radiology was -.29. It appeared that the Radiation Physics and Radiation Safety portions tested for on National Boards may not have been emphasized in the Problem-based Curriculum to the proper degree.

In Podiatric Medicine the effect size was -.08, an insignificant finding.

In Dermatology the effect size was .09, .11 for Surgery and .1 in Orthopaedics. The National Board Examinations are multiple-choice tests where recall or recognition of memorized facts are rewarded. The Problem-based Curriculum did not emphasize these facts, however, the problem-based students did score slightly better in these three sections than the students from the year before.

The overall effect size for the 6 sections was -.08, which was interpreted as insignificant. From this point of view, there did not seem to be a clear advantage to the use of problem-based learning based on the National Boards parameter.

Clinical Evaluations exhibited a higher effect size than did the National Boards. All were positive, indicating that the Class of 1991 did better on Clinical Evaluations than the Class of 1990. The effect sizes for Knowledge was .05 which indicated that the Class of 1990 had, for all intents and purposes, the same amount of knowledge as the Class of 1991. The effect size for Self-directedness was .07. It was hoped that this area would have a larger effect size considering the fact that problem-based learning should stimulate this characteristic due to the large amount of time students spend on their own discovering answers to problems (Grow, 1991).

The effect size for Skills was .19. Although problem-based learning does not in and of itself stimulate an increase in clinical skills, the Class of 1991, as part of the Problem-based Curriculum were administered an Objective Clinical Skills Examination before leaving the third year. This involved both problem-solving and the performance of certain clinical tasks. This was not done previously and may have resulted in increased confidence and readiness to perform these tasks in a true clinical setting.

The effect size for Interpersonal Skills was .56. This is completely understandable due to the fact that students were required to interact for up to 90 minutes daily in a group with a facilitator. This gave the students more confidence when placed in a hospital setting where, for the Class of 1991, the experience was not new. For the Class of 1990, this abrupt change from passive didactic learning to hospital setting may have resulted in a perceived introversion by evaluating clinicians. This could have been a primary reason the residency placement success was so high for the Class of 1991.

The overall effect size for Clinical Evaluations was .22. Although below the .33 that researchers feel was necessary for further investigation, the great success in the residency selection process may be a mitigating factor.

Residency program selection by the student is a multifactorial decision. Individual priorities and their balancing remain at the heart of the process (Simmonds, Robbins, Brinker, Rice & Kerstein, 1990). What is clear is that careers are made and broken on the selections of the residencey program. A student wants to be an Orthopaedic Surgeon and soon realizes that only one of five students that apply for

Orthopaedic Residencies is accepted. Years of school and thousands of dollars in tuition have been spent. If a student is not accepted into the desired program, he or she perceives that much may have been wasted and years of potential frustration lie ahead.

The Problem-based Curriculum appears to be responsible for the statistically significant success of the Class of 1991 in securing desired residencies. Residency selections are often based on an interview that includes a "work up" of a real case. The student is given a patient situation and invited to proceed with the next course of thought or action. This is a recreation of the Problem-based Curriculum where the same scenario is played out daily. The residency interview format, while new for some is a situation the problem-based learner has found him or herself in numerous times. The comfort level is much greater; hence the problem solving process greatly facilitated. This is communicated to the interviewer.

Recommendations For Action: From a practical point of view, 70% of the Class of 1991 was chosen by their first choice hospitals while only 42% of the Class of 1990 were chosen by their first choice hospitals. This great improvement would seem to necessitate that the Problem-based Curriculum be continued.

Two areas that appeared to be lacking in success were the Community Health and Radiology sections of the National Boards. Because the Jurisprudence portion of the Community Health section was not taught via the problem-based method, attempts should be made to insert this into the Problem-based Curriculum. The same should hold true for Radiation Physics and Radiation Safety. Problems could be written that would introduce the student to medical malpractice,

workman's compensation claims, expert witnessing, the impaired physician and the tort system. Problems could also be constructed that dealt directly with pregnant patients requiring X-ray examination, X-ray machines with radiation leaks, shielding and overexposure besides simply interpreting radiographs.

Recommendations for Further Research: One area of great interest should be measuring the efficiency of problem-based learning. After start up time investments, how much time does the faculty invest each year in problem-based learning? This should be compared to the time spent lecturing during the last year of the conventional curriculum.

The outcomes evaluated in this study were short term. Considerable effort should be spent looking at long term outcomes that may reflect on the Problem-based Curriculum. Factors such as Board Certification, publishing in refereed journals and teaching should be investigated as the potential for these career milestones are encountered.

Investigations should be made into the self-directedness of problem-based learners. Self-directedness is much like self-actualization (Long & Ashford, 1976). One is not self-actualized but self-actualizing is an ongoing process. Houle (1980) stated that learning for the professional begins in earnest after entry into practice. Graduates of both programs could be examined to determine if this occurs sooner with the physicians who encountered a problem-based curriculum.

Lastly, the long term goal of affecting individuals so that they become good physicians should be assessed in relation to both the Traditional and Problem-based Curricula. Sade, Stroud, Levine and

Fleming (1985) investigated the qualities of the superior physician. They found that the ability to convert acquired information into working knowledge ranked first. This was followed by keeping abreast of progress in medical knowledge and possessing the knowledge and the ability to study patients thoroughly. Emotional stability ranked fourth and having good doctor-patient relationships was fifth. The curriculum is a means to this end. The best curriculum to achieve this goal is the one that should be ultimately implemented.

Summary

A new curriculum was introduced into the College of Podiatric Medicine and Surgery at the University of Osteopathic Medicine and Health Sciences, Des Moines, Iowa. It was problem-based and replaced the third year clinical sciences that were traditionally taught through the lecture method. Small groups of approximately 7 students met almost daily with a facilitator who guided the students through real patient problems without giving them answers to questions. Students were tested by essay examination. At the end of their third year, students left the University for their fourth year to function as externs in hospital settings across the United States.

In order to evaluate the effectiveness of this new program, three parameters were examined. They were (1) National Board Scores, (2) Clinical Evaluations by physicians responsible for the students who were externs during their fourth year of school, and (3) Successful acceptance of the students into the residencies they desired most. The last class encountering a traditional curriculum, 1990, was compared to the first class to encounter the Problem-based Curriculum, the Class of 1991.

Statistically significant results were found in the Community Health portion of the National Boards where the Traditional students scored higher than the problem-based students ($p=.036$). On the Interpersonal Skills portion of the Clinical Evaluations, the problem-based students scored significantly better ($p=.003$) than the traditional students. The Class of 1991 had a statistically better ($p<.05$) acceptance into desired residencies than the Class of 1990 which had a traditional curriculum.

References

Baca, E., Mennin, S.P. Kaufman, A., & Moore-West, M. (1987). A comparison between a problem-based, community-oriented track and a traditional track within one medical school. InKhattab, T. (Ed.), Innovations in medical education; an evaluation of its present status. New York: Springer.

Barrows,H.S., & Abrahamson, S. (1964). The programmed patient: A technique for appraising student performance in clinical neurology. Journal of Medical Education, 39, 802-805.

Barrows, H.S., & Tamblyn, R.M. (1977). The portable patient problem pack (p_4): A problem-based learning unit. Journal of Medical Education, 52, 1002-4.

Barrows, H.S., & Tamblyn, R.M. (1980). Problem-based learning: An approach to medical education. New York: Springer.

Borg, W.A., & Gall, M.D. (1989). Educational research introduction. New York: Longman.

Bruner, J.S. (1961). The act of discovery. Harvard Educational Review, 31(1), 21-32.

Bugge, K.E. (1965). Skolen for livet: studier over n.f.s. grundtvigs paedagogiske tanker (The school for life: studies on n.f.s. grundtvig's pedagogical thoughts). Copenhagen: Institut for Dansk Kirkehistorie.

Chamberlin, T.C. (1890). The method of multiple working hypothesis. Science, 15, 92-97.

Claesen, H.F., & Boshuizen, H.P. (1985). Recall of medical information by students and doctors. Medical Education, 19, 61-67.

Coulson, R.L. (1983). Problem-based student-centered learning of the cardiovascular system using the problem-based learning module. (p.b.l.m.) The Physiologist, 26, 220-224.

Dewey, J. (1933). How we think: A restatement of the relation of reflective thinking to the educational process. Boston: Heath.

Eisenstaedt, R.S., Barry, W.E., & Glanz, K. (1990). Problem-based learning: cognitive retention and cohort traits of randomly selected participants and decliners. Academic Medicine, 65, (9) Sept. Supplement, S11-S12.

Elstein, A.S., & Bordage, G. (1979). Psychology of clinical reasoning. In G. Stone, F. Cohen, and N. Adler (Eds), Health psychology. San Francisco: Jossey-Bass.

Flexner, A. (1910). Medical education in the United States and Canada, A report to the Carnegie Foundation for the Advancement of Teaching , Bulletin No. 4. Boston: Updyke.

Gagne, R.M. (1977). The conditions of learning. New York: Holt Rinehart and Winston.

Glass, G.V., McGaw, B., & Smith, M.L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

Gonnella, J.S., & Hojat, M. (1983). Relationship between performance in medical school and postgraduate competence. Journal of Medical Education, 58, 679-685.

Grow, G.O. (1991). Teaching learners to be self-directed. Adult Education Quarterly, 41,(3), 125-149.

Grundtvig, N.F.S. (1976). Selected writings: N.F.S. Grundtvig. J. Knudsen, (Ed. and Trans.). Philadelphia: Fortress.

Ham, T.A. (1962). Medical education at Western Reserve University. The New England Journal of Medicine, 267, 868-874.

Ham, T.A. (1964). The training of the physician research in the teaching-learning process-two experimental modules. The New England Journal of Medicine, 271,1042-1046.

Harris, J.W., Horrigan,D.L., Ginther,J.R. & Ham, T.H. (1962) Pilot study in teaching hematology with emphasis on self-education by the students. Journal of Medical Education, 37,719-728.

Herber, H.C., & Nelson-Herber, J. (1987). Developing independent learners. Journal of Reading, 10, 584-588.

Houle, C.O. (1980). Continuing learning in the professions. San Francisco: Jossey-Bass.

Jones, J.W., Bieber, L.L., Echt, R., Scheifly, V., & Ways, P.O. (1984). A problem-based curriculum-10 years experience. In Schmidt, H.G., DeVolder,M.L. (Eds.). Tutorial in problem-based learning, Assen/Maastricht: Van Gorcum.

Katona, G. (1940). Organizing and memorizing. New York: Columbia University Press.

Kaufman, A., Menin, S., Waterman, R., Duban, S., Hansbarger, C., Silverblatt, H., Obenshain, S. S., Kantrowitz, M., Becker, T., Samet, J., & Wiese, W. (1989). The New Mexico experiment: Educational innovation and institutional change. Academic Medicine, 64, 285-294.

Knowles, M.S. (1980). The modern practice of adult education. New York: Cambridge.

Knowles, M.S, (1984). Andragogy in action, San Francisco: Jossey-Bass.

Lindeman, E.C. (1961). The meaning of adult education. Montreal: Harvest House.

Long, H.B., & Ashford, M.L. (1976). Self-directed inquiry as a method of continuing education in colonial america. The Journal of General Education, 28, 245-255.

McGuire, C.H. (1972). Research on identification of student needs Proceedings of the Fourth Panamerican Conference of Medical Education, Toronto, Ontario, Canada, pp. 154-160.

Miller, G.A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. The PsychologicalReview. 63, 81-97.

Miller, G.E. (1962). An inquiry into medical teaching. Journal of medical education, 37, 185-191.

Miller, G.E. (1978). The contributions of research in the learning process. Medical Education, 12, 28.

Neufeld,V.R., Woodward,C.A., & MacLeod,S.M. (1989). The McMaster m.d. program: A case study of renewal in medical education. Academic Medicine, 64, 423-432.

Newble, D.I., & Clarke,R.M. (1986). The approaches to learning of students in traditional and in an innovative problem-based medical school. Medical Education, 20, 267-273.

Tonai, A.K., Golmon,M.E., Shanks,C.A., Schafer,M.F., & Brunner,E.A.(1984). Relationship between past academic performance and results of specialty in training examinations. Journal of Medical Education, 59, 341-344.

Sade, R.M., Stroud, M.R., Levine, J.H., & Fleming, G.A. (1985). Criteria for selection of future physicians. Annals of Surgery, 201(2), 225-230.

Saunders, N.A. (1987). A programme of outcome evaluation studies: experiences at the university of new castle. In Khattab, T. Schmidt, H.G., Nooman, Z, and Ezzat, E. (Eds.). Innovations in medical education: An evaluation of its present status. . New York: Springer.

Simmonds, A.C., Robbins, J.M., Brinker, M.R., Rice, J.C., & Kerstein, M.D. (1990). Factors important to students in selecting a residency program. Academic Medicine, 65, 640-643.

Stewart, D.W. (1987). Adult learning in America: Eduard Lindeman and his agenda for lifelong education. Malabar Florida: Krieger.

Tomczak, R.L. & Dowdy, N.L. (1988). Selection of podiatric surgery residents. Journal of Foot Surgery, 27, 240-242.

Veloski, J., Herman, M.W., Gonnella, J.S. Zeleznik,C., & Kellow, W.F. (1979). Relationships between performance in medical school and first postgraduate year. Journal of Medical Education, 54, 909-916.

Wearn, J.T. (1956). Western reserve 1. background and philosophy of experiment. Journal of Medical Education, 31, 516-523.

West, K.M. (1966) The case against teaching. Journal of Medical Education, 41, 766-771.

Williams, G. (1980). Western Reserve's experiment in medical education and its outcome. New York: Oxford University Press.

Williams,T., Sachs,L., & Veloski, J.J. (1986). Performance on the NBME part II examination and career choice. Journal of Medical Education, 61, 979-981.

Woodward, C.A. (1984). Summary of McMaster graduates performance on the medical council; of Canada examination. Hamilton, Ontario,Canada: McMaster University Faculty of Health Sciences.

Woodward,C.A., McCauley,R.G., & Ridge,H. (1981). Unraveling the meaning of global comparative ratings of interns. In Research in Medical Education. Washington, DC: Association of American Medical Colleges.

Worthen, B.R. (1968). Discovery and expository task presentation in elementary mathematics. Journal of Educational Psychology Monograph Supplement, 59, No.1, Part 2.